

# Optimasi Prediksi Produksi Padi & Beras Sumatera via Big Data Pipeline

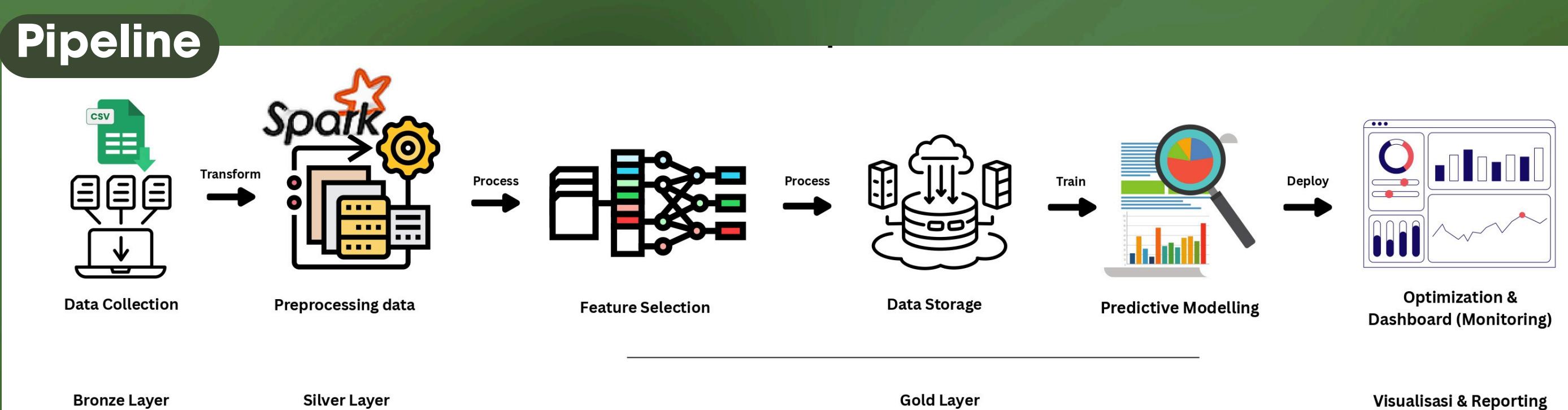


## Definisi

Pulau Sumatera berperan strategis sebagai lumbung padi dan beras nasional, namun menghadapi tantangan variabilitas produksi yang memerlukan sistem peramalan dan terintegrasi. Oleh karena itu, dibutuhkan pendekatan teknologi berbasis big data yang mampu menyediakan prediksi produksi untuk mendukung perencanaan pertanian dan kebijakan distribusi pangan yang lebih efektif.

## Alur Pemrosesan

1. Ingesti Data, mengumpulkan data produksi padi dan disimpan ke sistem HDFS.
2. Transformasi, *preprocessing* dan simpan dalam hive dan spark.
3. *Processing*, melatih serta evaluasi model (*Random Forest Regression*) dengan Spark MLlib.
4. Visualisasi, membuat *dashboard* dengan Power BI
5. Mendapatkan model untuk dilakukan implementasi

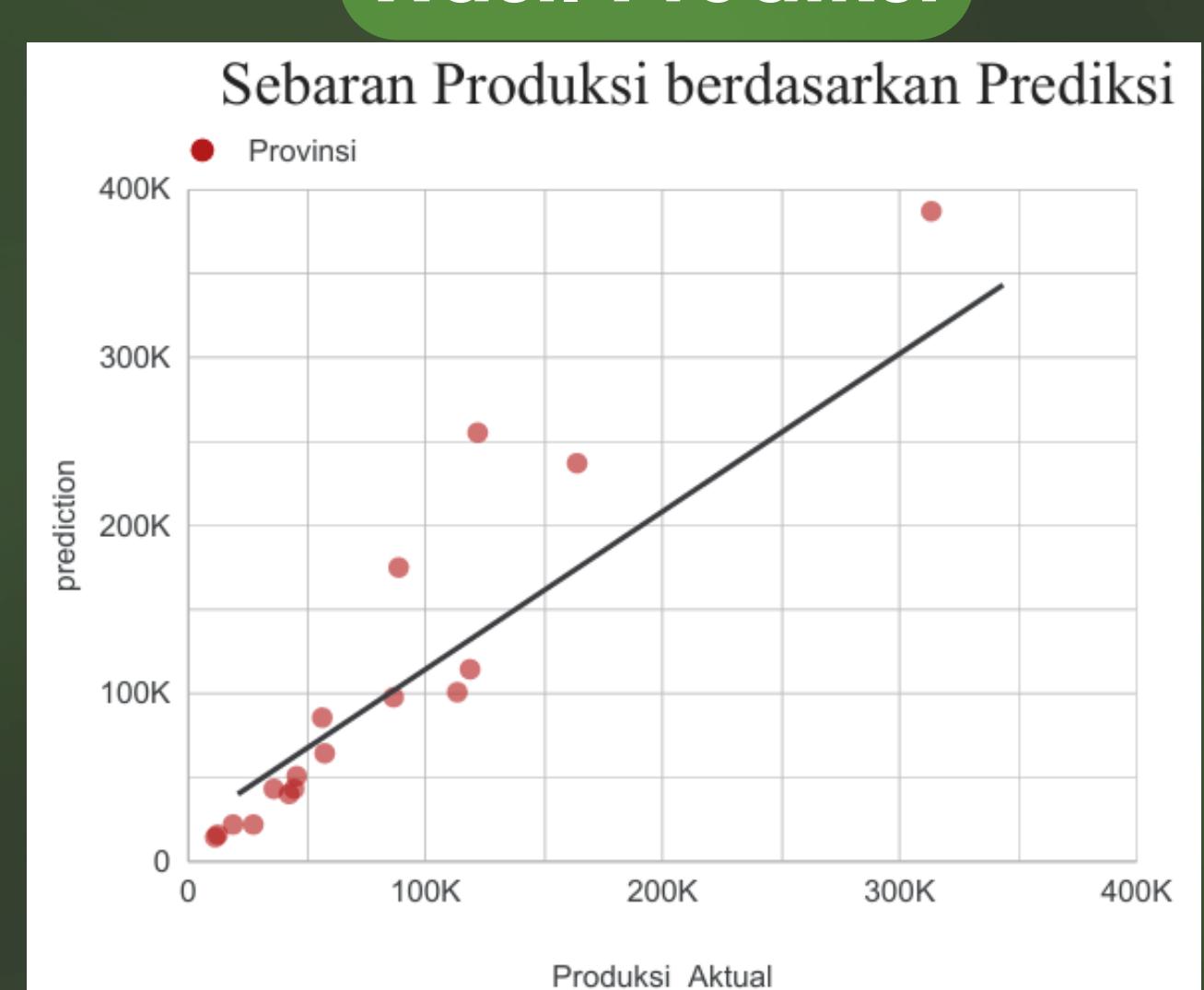


## Hasil Prediksi

Hasil visualisasi menjelaskan bahwa warna hijau memiliki nilai prediksi yang mendekati dengan nilai aktualnya, hal ini menandakan bahwa model dengan baik menjelaskan data.



## Hasil Prediksi



## Evaluasi Model

